2020-04-10 Cornell LD4P2 Meeting notes

Date: 10 Apr 2020

Attendees: Jason, Steven, Tim, John, Lynette, Huda, Simeon

Regrets:

Agenda & Notes

Review actions from 2020-04-03 Cornell LD4P2 Meeting notes

- E. Lynette Rayle QA performance
  - 2020-03-06: Dave in process of converting everything over. Unsure of status for any authority, including MeSH. When LC done, we'll know whether this has impact since there are considerable usage data for LC. Dave and Lynette each working on other projects at the moment.
  - 2020-03-20: Dave is tied up with dept. COVID-19 planning. Log analysis shows two classes of error: 1) 502 is Dave's cache not responding, 2) 500 error which depends on query, Java null pointer error from Dave's system, and fewer of a third 3) timeout error
  - 2020-04-03 On hold because Dave is otherwise tied up
  - 2020-04-10 Lynette has done an analysis of current performance Performance Analysis for QA Lookups – The Getty authorities are pretty much all sub-second which is not too bad, however averaging over all authorities we are around 2s. We find that for ARGOVOC and GEONAMES direct access (without extended context) but when going against the cache we find typically 3-4s with the QA portion (graph load and normalization) approaching a second – Lynette is not yet sure why Geonames has such a large normalization hit for example, why is the graph load time long for Argovoc?
- E. Lynette Rayle to ask Tiziana about SHARE-VDE APIs for real-time up-to-date search and for possible engagement in linked data best practices for authoritative data working group
  - 2020-03-20: Will instead email a larger set of people (individually) about proposed best practices working group, will delay a bit until "new normal" has set in
  - 2020-04-10 Got positive responses from some folks about participating including LoC and Casalini, maybe from OCLC and still following up with others documenting on Charter 1 - Best Practices for Authoritative Data Working Group, with target start in early May
- Simeon Warner to ask PIs and Michelle about planning for partner meeting
  - 2020-04-03 Ongoing discussion about a set of 1-2h calls for partner meeting

Status updates and planning

- Enhanced Discovery - **WHAM!** (see also https://wiki.duraspace.org/x/sJI7Bg and https://github.com/LD4P/discovery/projects/1)
  - https://docs.google.com/document/d/1u88qVOhp92C6Y1N9dNpNtjV_XOnusFB0nb2aJ-Q10FM/edit
  - 2020-03-20 Consensus from discussion that top priority should be work on the autosuggest, with possibility of additional work on discogs or syllabus. Goal for the next week is to develop plan for autosuggest work and to understand any possible issues
  - 2020-03-25 Discussion with developers, Notes, Slides
  - 2020-04-03 Organizing doc
  - John notes that we already have a "best bets" autocomplete on our live catalog search (e.g. nature for Nature Journal or hours for library hours), this was previously disabled in our tests but we likely need some way to preserve best bets along with an enhanced autocomplete. Tim suggests that best bets might be a special category that shows top of the list
  - Tim working on getting real data in, Huda working on query matchings (including issues with period etc.), get main index set ip with data including variants, John continuing to look at packaging
  - Later - think of pseudonyms from wikidata, maybe VIAF
- Cataloging Sinatra and other 45's (Discogs data, https://github.com/ld4p/qa_server/issues?q=is%3Aissue+is%3Aopen+label%3ADiscogs)
  - Lookups for place not usable and hence places are not being recorded, relies on work from Dave to fix: https://github.com/LD4P/qa_server/issues/248 & https://github.com/LD4P/qa_server/issues/240
  - Have a currently insurmountable issue with nested profiles. When create Work profile with nested Instance profile there isn't a URI for the Instance (it just gets hung from a bnode). Without a URI the title of the Instance doesn't get indexed. The Sinopia team are unable to fix this in the near term.
  - 2020-03-20 not expecting progress in the near future
- Authority Lookups for Sinopia (Lookup infrastructure: https://github.com/LD4P/qa_server/projects/2, Authority requests: https://github.com/LD4P/qa_server/projects/1)
  - 2020-03-27 A couple of updates being put into production now: ISNI now available, Wikidata wasn't working properly because of a lack of response header which has now been added so that it plays nicely with Sinopia, also fix with CERL. Doing some infrastructure work to use Reddis for caching and avoid a memory leak issue, working with Greg to understand issues
  - 2020-04-03 Recently done: cleaned up and updated UBER Issue for Authorities (#253), UBER Issue for Indexing (#239), OCLC FAST cache vs. direct (Sinopia #2125), Use Case: Authority Lookup UI option in place of auto-complete (Sinopia #2123)
- Travel and meetings (see LD4P2 Cornell Meeting Attendances)
  - LD4P2 cohort meeting, LC, previously scheduled 20
    - Current discussions that Michelle has organized are reviewing status of cohort and whether some remote meeting is needed – lots of COVID-19 discussion and changes to plans
  - LD4P2 remote partner meetings
    - Sinopia topics
      - Possible need for lookups and not just auto-complete to meet cataloger needs: https://github.com/LD4P/sinopia_editor/issues/2123
  - Knowledge Graph Conference (Columbia University), May 6-7, workshops 5/4-5/5
    - Now remote and Huda confirmed, sticking with original dates, don't yet know scheduling
  - LD4 Conference 2020 at College Station, TX (TAMU) - May 13/14, 2020
    - Planning going ahead for meeting to work out virtual presentations. There was a survey of presenters to understand how they might want to present. Steven likes the idea of recorded sessions followed by office hours.
  - DCMI (9/14-9/17, Ottawa): CFP due 4/13
    - Not cancelled

- - rdfs:seeAlso [Conferences Related to Linked Data in Libraries](#)
- Next meetings
  - 2020-04-17 will do 9-10am to avoid conflict with FOLIO Friday